```
GATCACACCC CTCCCACCCT TCTCTTTTCA AGGTTGTCCC CTTCTCCCAC 50
GGCTTTATGT ACTTCCCACT CTMTAATTCG CTCTTTCCAT TCCAAGCCAA 100
GCAACATCTG TGAGCAGCTC ATCCTTCCCA ATATGGGCGA ATGGCAGGAG 150
CAGATGATGG GTTTTGACGT GGAGGATGTT CTGTCTCAGC TGAGCCAAAA 200
TGAGAAGATT GCTCTCTTGT CCGGCATTGA TTTCTGGCAT ACTTATCCCA 250
TACCAAAGTA CAACGTCCCT TCAGTCCGCC TAACGGACGG TCCTAACGGC 300
ATACGAGGCA CAAAGTTTTT TGCTGGCATT CCTGCTGCCT GCCTGCCATG 350
TGGGACGGCC CTGGCCTCTA CCTGGGATAA GCAGCTGCTG AAGAAGGCTG 400
GGAAGCTGCT CGGTGATGAG TGCATCGCAA AAGGCGCCCA CTGCTGGCTG 450
GGCCCAACAA TCAATACTCC CCGATCTCCT CTGGGGGGGC GCGGCTTCGA 500
GTCATTTTCG GAAGATCCGT ACCTGTCCGG CATCCTTGCT GCATCTATGA 550
TTCTCGGCTG TGAAAGCACA GGTGTCATCT CTGCCGTCAA ACACTTTGTC 600
GCCAACGACC AGGAGCACGA GCGGCGAGCG GTCGACTGTC TCATCACCCA 650
GCGGGCTCTC CGGGAGGTCT ATCTGCGACC CTTCCAGATC GTAGCCCGAG 700
ATGCAAGGCC CGGCGCATTG ATGACATCCT ACAACAAGGT CAATGGCAAG 750
CACGTCGCTG ACAGCGCCGA GTTCCTTCAG GGCATTCTCC GGACTGAGTG 800
GAATTGGGAT CCTCTCATTG TCAGCGACTG GTACGGCACC TACACCACTA 850
TTGATGCCAT CAAAGCCGGC CTTGATCTCG AGATGCCGGG CGTTTCACGA 900
TATCGCGGCA AATACATCGA GTCTGCTCTG CAGGCCCGTT TGCTGAAGCA 950
GTCCACTATC GATGAGCGCG CTCGCCGCGT GCTCAGGTTC GCCCAGAAGG 1000
CCAGCCATCT CAAGGTCTCC GAGGTAGAGC AAGGCCGTGA CTTCCCAGAG 1050
GATCGCGTCC TCAACCGTCA GATCTGCGGC AGCAGCATTG TCCTACTGAA 1100
GAATGAGAAC TCCATCTTAC CTCTCCCCAA GTCCGTCAAG AAGGTCGCCC 1150
TTGTTGGATC CCACGTGCGT CTACCGGCTA TCTCGGGAGG AGGCAGCGCC 1200
TCTCTTGTCC CTTACTATGC CATATCTCTA TACGATGCCG TCTCTGAGGT 1250
ACTAGCCGGT GCCACGATCA CGCACGAGGT CGGTGCCTAT GCCCACCAAA 1300
TGCTGCCCGT CATCGACGCA ATGATCAGCA ACGCCGTAAT CCACTTCTAC 1350
AACGACCCCA TCGATGTCAA AGACAGAAAG CTCCTTGGCA GTGAGAACGT 1400
ATCGTCGACA TCGTTCCAGC TCATGGATTA CAACAACATC CCAACGCTCA 1450
ACAAGGCCAT GTTCTGGGGT ACTCTCGTGG GCGAGTTTAT CCCTACCGCC 1500
ACGGGAATTT GGGAATTTGG CCTCAGTGTC TTTGGCACTG CCGACCTTTA 1550
TATTGATAAT GAGCTCGTGA TTGAAAATAC AACACATCAG ACGCGTGGTA 1600
CCGCCTTTTT CGGAAAGGGA ACGACGGAAA AAGTCGCTAC CAGGAGGATG 1650
GTGGCCGGCA GCACCTACAA GCTGCGTCTC GAGTTTGGGT CTGCCAACAC 1700
GACCAAGATG GAGACGACCG GTGTTGTCAA CTTTGGCGGC GGTGCCGTAC 1750
ACCTGGGTGC CTGTCTCAAG GTCGACCCAC AGGAGATGAT TGCGCGGGCC 1800
GTCAAGGCCG CAGCCGATGC CGACTACACC ATCATCTGCA CGGGACTCAG 1850
CGGCGAGTGG GAGTCTGAGG GTTTTGACCG GCCTCACATG GACCTGCCCC 1900
```

FIG._1A

```
CTGGTGTGGA CACCATGATC TCGCAAGTTC TTGACGCCGC TCCCAATGCT 1950
GTAGTCGTCA ACCAGTCAGG CACCCCAGTG ACAATGAGCT GGGCTCATAA 2000
AGCAAAGGCC ATTGTGCAGG CTTGGTATGG TGGTAACGAG ACAGGCCACG 2050
GAATCTCCGA TGTGCTCTTT GGCAACGTCA ACCCGTCGGG GAAACTCTCC 2100
CTATCGTGGC CAGTCGATGT GAAGCACAAC CCAGCATATC TCAACTACGC 2150
CAGCGTTGGT GGACGGGTCT TGTATGGCGA GGATGTTTAC GTTGGCTACA 2200
AGTTCTACGA CAAAACGGAG AGGGAGGTTC TGTTTCCTTT TGGGCATGGC 2250
CTGTCTTACG CTACCTTCAA GCTCCCAGAT TCTACCGTGA GGACGGTCCC 2300
CGAAACCTTC CACCCGGACC AGCCCACAGT AGCCATTGTC AAGATCAAGA 2350
ACACGAGCAG TGTCCCGGGC GCCCAGGTCC TGCAGCTATA CATTTCGGCC 2400
CCAAACTCGC CTACACATCG CCCGGTCAAG GAGCTGCACG GATTCGAAAA 2450
GGTGTATCTT GAAGCTGGCG AGGAGAAGGA GGTACAAATA CCCATTGACC 2500
AGTACGCTAC TAGCTTCTGG GACGAGATTG AGAGCATGTG GAAGAGCGAG 2550
AGGGGCATTT ATGATGTGCT TGTAGGATTC TCGAGTCAGG AAATCTCGGG 2600
CAAGGGGAAG CTGATTGTGC CTGAAACGCG ATTCTGGATG GGGCTGTAGA 2650
TTCAACACGT GAGCAAAAGC GATTGCGGAA AGTACCAGAA AAGCCAAGGG 2700
AGTCAAAGGA TGGGAACTTG TGTCAATAGA AGATATGCAT GATGGGCATT 2750
TGGGATGGTG TTTGGCATTA TGCAAAGAAG CAAAGATGGA GTGATAAAAA 2800
AAAAAAAAAA AA                                          2812
```

FIG. 1B

```
MGEWQEQMMG FDVEDVLSQL SQNEKIALLS GIDFWHTYPI PKYNVPSVRL      50
TDGPNGIRGT KFFAGIPAAC LPCGTALAST WDKQLLKKAG KLLGDECIAK     100
GAHCWLGPTI NTPRSPLGGR GFESFSEDPY LSGILAASMI LGCESTGVIS     150
AVKHFVANDQ EHERRAVDCL ITQRALREVY LRPFQIVARD ARPGALMTSY     200
NKVNGKHVAD SAEFLQGILR TEWNWDPLIV SDWYGTYTTI DAIKAGLDLE     250
MPGVSRYRGK YIESALQARL LKQSTIDERA RRVLRFAQKA SHLKVSEVEQ     300
GRDFPEDRVL NRQICGSSIV LLKNENSILP LPKSVKKVAL VGSHVRLPAI     350
SGGGSASLVP YYAISLYDAV SEVLAGATIT HEVGAYAHQM LPVIDAMISN     400
AVIHFYNDPI DVKDRKLLGS ENVSSTSFQL MDYNNIPTLN KAMFWGTLVG     450
EFIPTATGIW EFGLSVFGTA DLYIDNELVI ENTTHQTRGT AFFGKGTTEK     500
VATRRMVAGS TYKLRLEFGS ANTTKMETTG VVNFGGGAVH LGACLKVDPQ     550
EMIARAVKAA ADADYTIICT GLSGEWESEG FDRPHMDLPP GVDTMISQVL     600
DAAPNAVVVN QSGTPVTMSW AHKAKAIVQA WYGGNETGHG ISDVLFGNVN     650
PSGKLSLSWP VDVKHNPAYL NYASVGGRVL YGEDVYVGYK FYDKTEREVL     700
FPFGHGLSYA TFKLPDSTVR TVPETFHPDQ PTVAIVKIKN TSSVPGAQVL     750
QLYISAPNSP THRPVKELHG FEKVYLEAGE EKEVQIPIDQ YATSFWDEIE     800
SMWKSERGIY DVLVGFSSQE ISGKGKLIVP ETRFWMGL                  838
```

Figure 2

```
MMGFDVEDVL SQLSQNEKIA LLSGIDFWHT YPIPKYNVPS VRLTDGPNGI      50
RGTKFFAGIP AACLPCGTAL ASTWDKQLLK KAGKLLGDEC IAKGAHCWLG     100
PTINTPRSPL GGRGFESFSE DPYLSGILAA SMILGCESTG VISAVKHFVA     150
NDQEHERRAV DCLITQRALR EVYLRPFQIV ARDARPGALM TSYNKVNGKH     200
VADSAEFLQG ILRTEWNWDP LIVSDWYGTY TTIDAIKAGL DLEMPGVSRY     250
RGKYIESALQ ARLLKQSTID ERARRVLRFA QKASHLKVSE VEQGRDFPED     300
RVLNRQICGS SIVLLKNENS ILPLPKSVKK VALVGSHVRL PAISGGGSAS     350
LVPYYAISLY DAVSEVLAGA TITHEVGAYA HQMLPVIDAM ISNAVIHFYN     400
DPIDVKDRKL LGSENVSSTS FQLMDYNNIP TLNKAMFWGT LVGEFIPTAT     450
GIWEFGLSVF GTADLYIDNE LVIENTTHQT RGTAFFGKGT TEKVATRRMV     500
AGSTYKLRLE FGSANTTKME TTGVVNFGGG AVHLGACLKV DPQEMIARAV     550
KAAADADYTI ICTGLSGEWE SEGFDRPHMD LPPGVDTMIS QVLDAAPNAV     600
VVNQSGTPVT MSWAHKAKAI VQAWYGGNET GHGISDVLFG NVNPSGKLSL     650
SWPVDVKHNP AYLNYASVGG RVLYGEDVYV GYKFYDKTER EVLFPFGHGL     700
SYATFKLPDS TVRTVPETFH PDQPTVAIVK IKNTSSVPGA QVLQLYISAP     750
NSPTHRPVKE LHGFEKVYLE AGEEKEVQIP IDQYATSFWD EIESMWKSER     800
GIYDVLVGFS SQEISGKGKL IVPETRFWMG L                          831
```

Figure 3

```
ATGGGCGAAT GGCAGGAGCA GATGATGGGT TTTGACGTGG AGGATGTTCT      50
GTCTCAGCTG AGCCAAAATG AGAAGATTGC TCTCTTGTCC GGCATTGATT     100
TCTGGCATAC TTATCCCATA CCAAAGTACA ACGTCCCTTC AGTCCGCCTA     150
ACGGACGGTC CTAACGGCAT ACGAGGCACA AAGTTTTTTG CTGGCATTCC     200
TGCTGCCTGC CTGCCATGTG GGACGGCCCT GGCCTCTACC TGGGATAAGC     250
AGCTGCTGAA GAAGGCTGGG AAGCTGCTCG GTGATGAGTG CATCGCAAAA     300
GGCGCCCACT GCTGGCTGGG CCCAACAATC AATACTCCCC GATCTCCTCT     350
GGGGGGGCGC GGCTTCGAGT CATTTTCGGA AGATCCGTAC CTGTCCGGCA     400
TCCTTGCTGC ATCTATGATT CTCGGCTGTG AAAGCACAGG TGTCATCTCT     450
GCCGTCAAAC ACTTTGTCGC CAACGACCAG GAGCACGAGC GGCGAGCGGT     500
CGACTGTCTC ATCACCCAGC GGGCTCTCCG GGAGGTCTAT CTGCGACCCT     550
TCCAGATCGT AGCCCGAGAT GCAAGGCCCG GCGCATTGAT GACATCCTAC     600
AACAAGGTCA ATGGCAAGCA CGTCGCTGAC AGCGCCGAGT TCCTTCAGGG     650
CATTCTCCGG ACTGAGTGGA ATTGGGATCC TCTCATTGTC AGCGACTGGT     700
ACGGCACCTA CACCACTATT GATGCCATCA AAGCCGGCCT TGATCTCGAG     750
ATGCCGGGCG TTTCACGATA TCGCGGCAAA TACATCGAGT CTGCTCTGCA     800
GGCCCGTTTG CTGAAGCAGT CCACTATCGA TGAGCGCGCT CGCCGCGTGC     850
TCAGGTTCGC CCAGAAGGCC AGCCATCTCA AGGTCTCCGA GGTAGAGCAA     900
GGCCGTGACT TCCCAGAGGA TCGCGTCCTC AACCGTCAGA TCTGCGGCAG     950
CAGCATTGTC CTACTGAAGA ATGAGAACTC CATCTTACCT CTCCCCAAGT    1000
CCGTCAAGAA GGTCGCCCTT GTTGGATCCC ACGTGCGTCT ACCGGCTATC    1050
TCGGGAGGAG GCAGCGCCTC TCTTGTCCCT TACTATGCCA TATCTCTATA    1100
CGATGCCGTC TCTGAGGTAC TAGCCGGTGC CACGATCACG CACGAGGTCG    1150
GTGCCTATGC CCACCAAATG CTGCCCGTCA TCGACGCAAT GATCAGCAAC    1200
GCCGTAATCC ACTTCTACAA CGACCCCATC GATGTCAAAG ACAGAAAGCT    1250
CCTTGGCAGT GAGAACGTAT CGTCGACATC GTTCCAGCTC ATGGATTACA    1300
ACAACATCCC AACGCTCAAC AAGGCCATGT TCTGGGGTAC TCTCGTGGGC    1350
GAGTTTATCC CTACCGCCAC GGGAATTTGG GAATTTGGCC TCAGTGTCTT    1400
TGGCACTGCC GACCTTTATA TTGATAATGA GCTCGTGATT GAAAATACAA    1450
CACATCAGAC GCGTGGTACC GCCTTTTTCG GAAAGGGAAC GACGGAAAAA    1500
GTCGCTACCA GGAGGATGGT GGCCGGCAGC ACCTACAAGC TGCGTCTCGA    1550
GTTTGGGTCT GCCAACACGA CCAAGATGGA GACGACCGGT GTTGTCAACT    1600
TTGGCGGCGG TGCCGTACAC CTGGGTGCCT GTCTCAAGGT CGACCCACAG    1650
GAGATGATTG CGCGGGCCGT CAAGGCCGCA GCCGATGCCG ACTACACCAT    1700
CATCTGCACG GGACTCAGCG GCGAGTGGGA GTCTGAGGGT TTTGACCGGC    1750
CTCACATGGA CCTGCCCCCT GGTGTGGACA CCATGATCTC GCAAGTTCTT    1800
GACGCCGCTC CCAATGCTGT AGTCGTCAAC CAGTCAGGCA CCCCAGTGAC    1850
AATGAGCTGG GCTCATAAAG CAAAGGCCAT TGTGCAGGCT TGGTATGGTG    1900
GTAACGAGAC AGGCCACGGA ATCTCCGATG TGCTCTTTGG CAACGTCAAC    1950
CCGTCGGGGA AACTCTCCCT ATCGTGGCCA GTCGATGTGA AGCACAACCC    2000
AGCATATCTC AACTACGCCA GCGTTGGTGG ACGGGTCTTG TATGGCGAGG    2050
```

FIG. 4A

```
ATGTTTACGT TGGCTACAAG TTCTACGACA AAACGGAGAG GGAGGTTCTG      2100
TTTCCTTTTG GGCATGGCCT GTCTTACGCT ACCTTCAAGC TCCCAGATTC      2150
TACCGTGAGG ACGGTCCCCG AAACCTTCCA CCCGGACCAG CCCACAGTAG      2200
CCATTGTCAA GATCAAGAAC ACGAGCAGTG TCCCGGGCGC CCAGGTCCTG      2250
CAGCTATACA TTTCGGCCCC AAACTCGCCT ACACATCGCC CGGTCAAGGA      2300
GCTGCACGGA TTCGAAAAGG TGTATCTTGA AGCTGGCGAG GAGAAGGAGG      2350
TACAAATACC CATTGACCAG TACGCTACTA GCTTCTGGGA CGAGATTGAG      2400
AGCATGTGGA AGAGCGAGAG GGGCATTTAT GATGTGCTTG TAGGATTCTC      2450
GAGTCAGGAA ATCTCGGGCA AGGGGAAGCT GATTGTGCCT GAAACGCGAT      2500
TCTGGATGGG GCTGTAG                                          2517
```

FIG._4B